CS623: Introduction to Computing with Neural Nets

Pushpak Bhattacharyya
Computer Science and Engineering Department
IIT Bombay

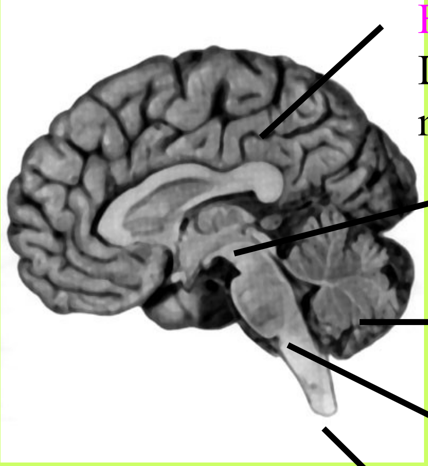
The human brain

Seat of consciousness and cognition

Perhaps the most complex information processing machine in nature

Historically, considered as a monolithic information processing machine

Beginner's Brain Map



Forebrain (Cerebral Cortex): Language, maths, sensation, movement, cognition, emotion

Midbrain: Information Routing; involuntary controls

Cerebellum: Motor Control

Hindbrain: Control of breathing, heartbeat, blood circulation

Spinal cord: Reflexes, information highways between body & brain

Brain : a computational machine?

Information processing: brains vs computers

- brains better at perception / cognition
- slower at numerical calculations
- parallel and distributed Processing
- associative memory

Brain : a computational machine? *(contd.)*

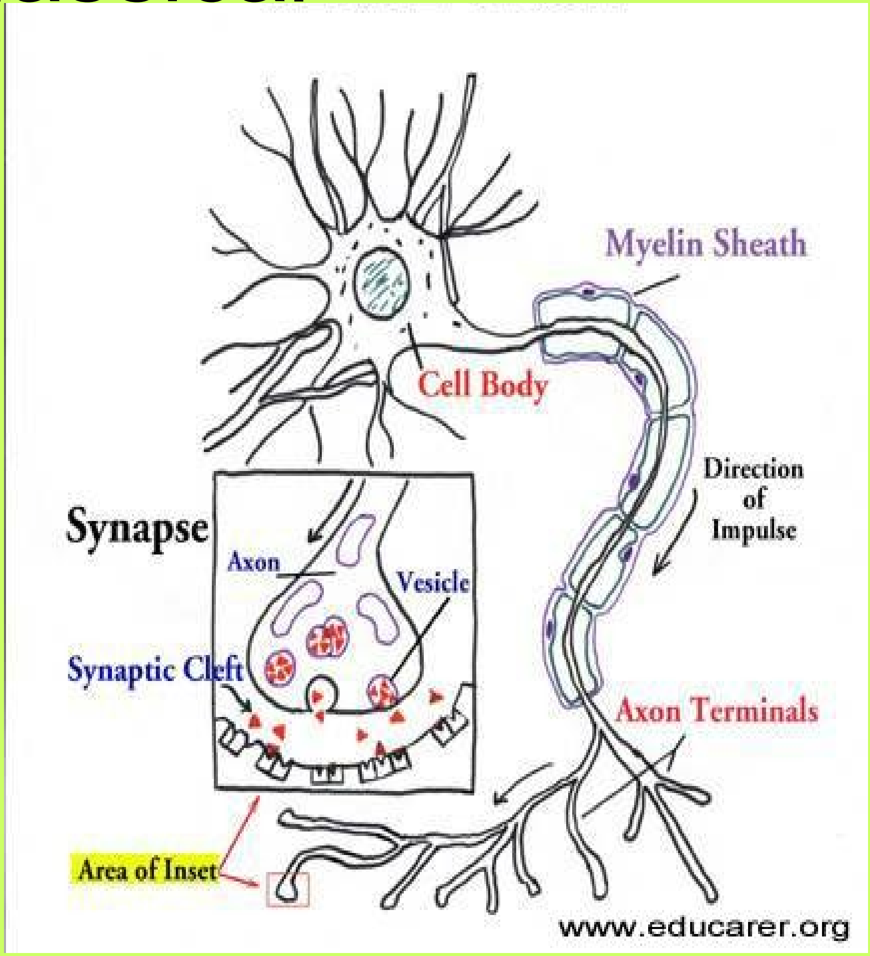
- Evolutionarily, brain has developed algorithms most suitable for survival
- Algorithms unknown: the search is on
- Brain astonishing in the amount of information it processes
  - Typical computers: 10^9 operations/sec
  - Housefly brain: 10^{11} operations/sec

Brain facts & figures

- Basic building block of nervous system: nerve cell (neuron)
- ~ 10^{12} neurons in brain
- ~ 10^{15} connections between them
- Connections made at "synapses"
- The speed: events on millisecond scale in neurons, nanosecond scale in silicon chips

Neuron - “classical”

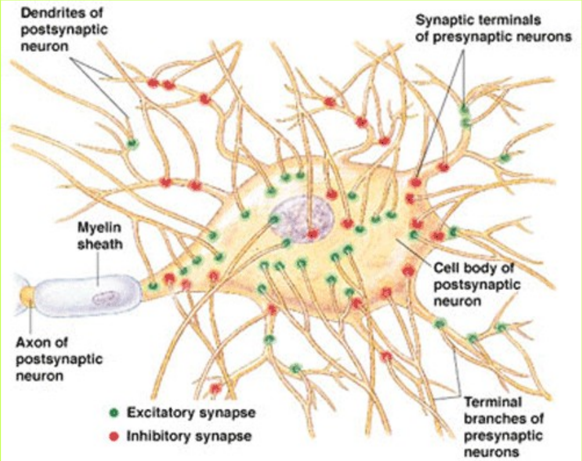
- Dendrites
  - Receiving stations of neurons
  - Don't generate action potentials
- Cell body
  - Site at which information received is integrated
- Axon
  - Generate and relay action potential
  - Terminal
    - Relays information to next neuron in the pathway



http://www.educarer.com/images/brain-nerve-axon.jpg

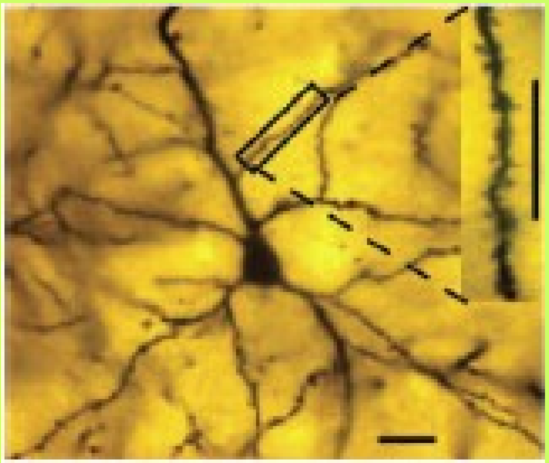
Computation in Biological Neuron

- Incoming signals from synapses are summed up at the soma
- Σ , the biological "inner product"
- On crossing a threshold, the cell "fires" generating an action potential in the axon hillock region

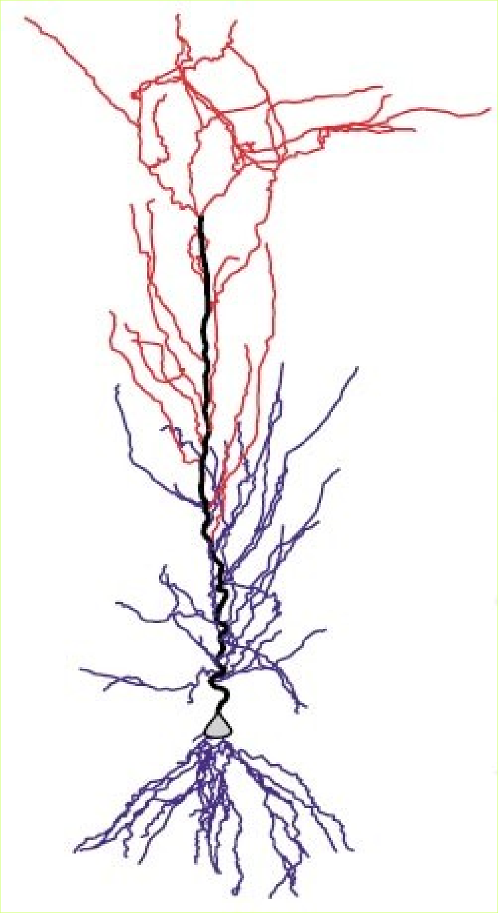


Synaptic inputs: Artist's conception

The biological neuron



Pyramidal neuron, from the amygdala (Rupshi *et al.* 2005)



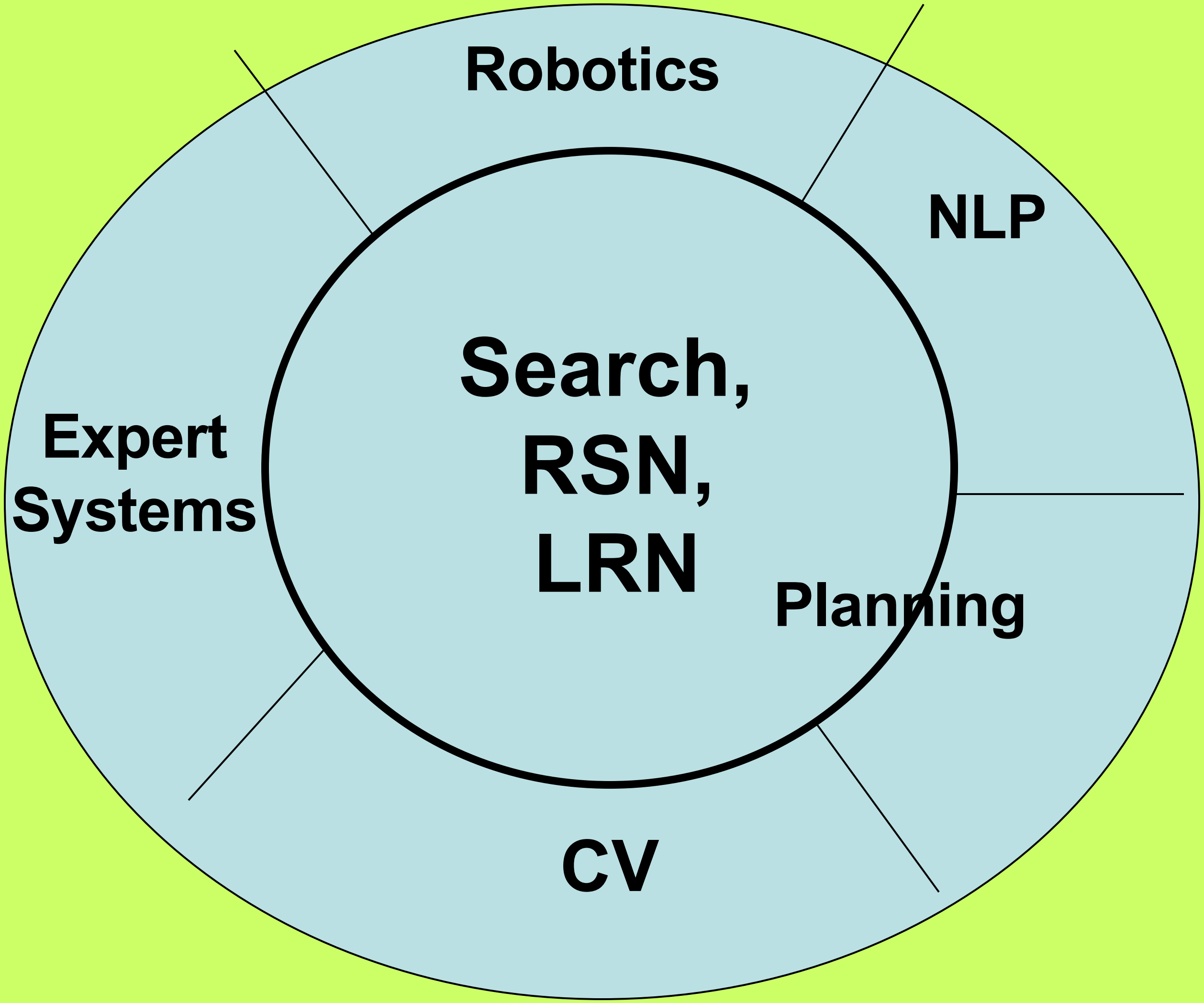
A CA1 pyramidal neuron (Mel *et al*. 2004)

A perspective of AI

Artificial Intelligence - Knowledge based computing

Disciplines which form the core of AI - inner circle

Fields which draw from these disciplines - outer circle.



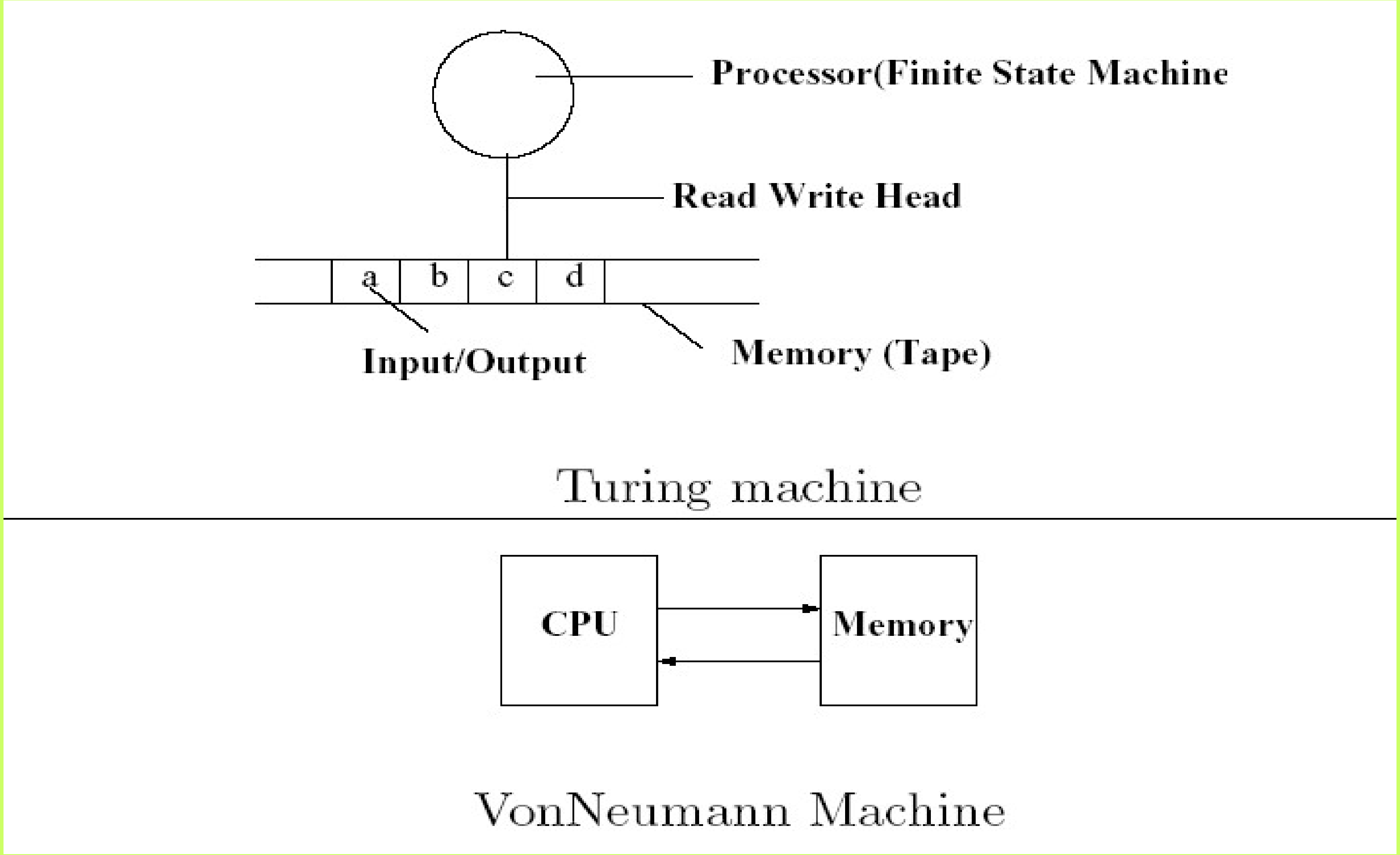
Symbolic AI

Connectionist AI is contrasted with Symbolic AI
Symbolic AI - Physical Symbol System Hypothesis
Every intelligent system can be constructed by storing and processing symbols and nothing more is necessary.

Symbolic AI has a bearing on models of computation such as

- Turing Machine
- Von Neumann Machine
- Lambda calculus

Turing Machine & Von Neumann Machine



Challenges to Symbolic AI

Motivation for challenging Symbolic AI

A large number of computations and information process tasks that living beings are comfortable with, are not performed well by computers!

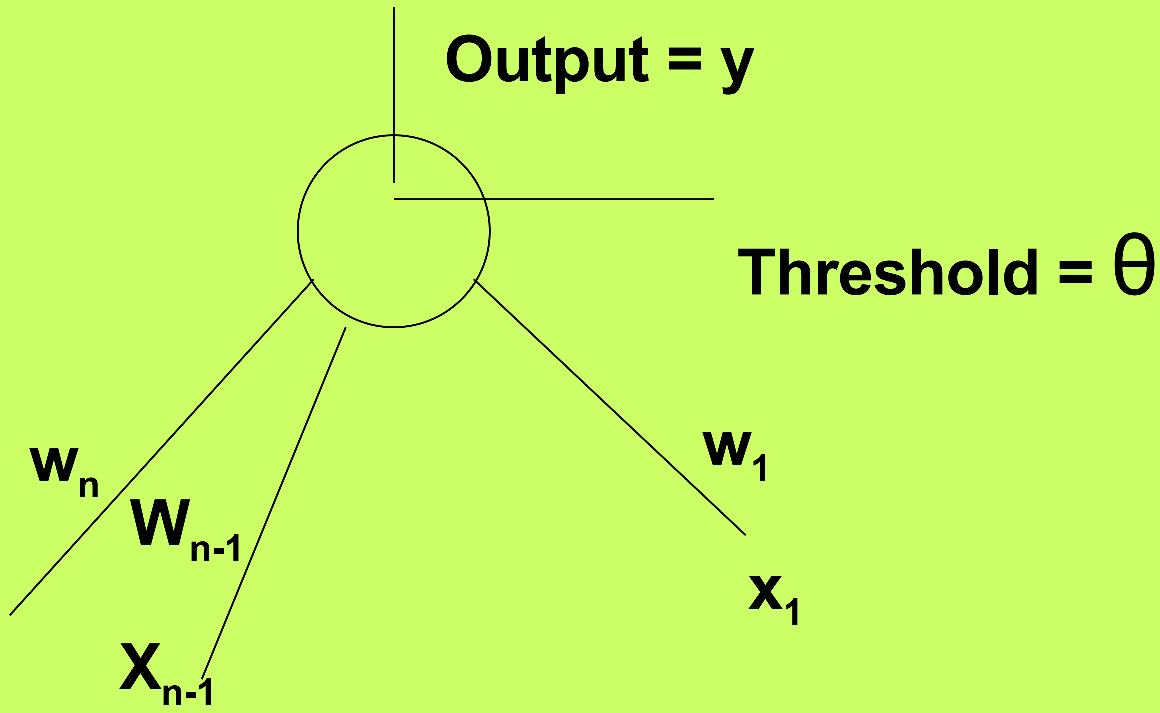
The Differences

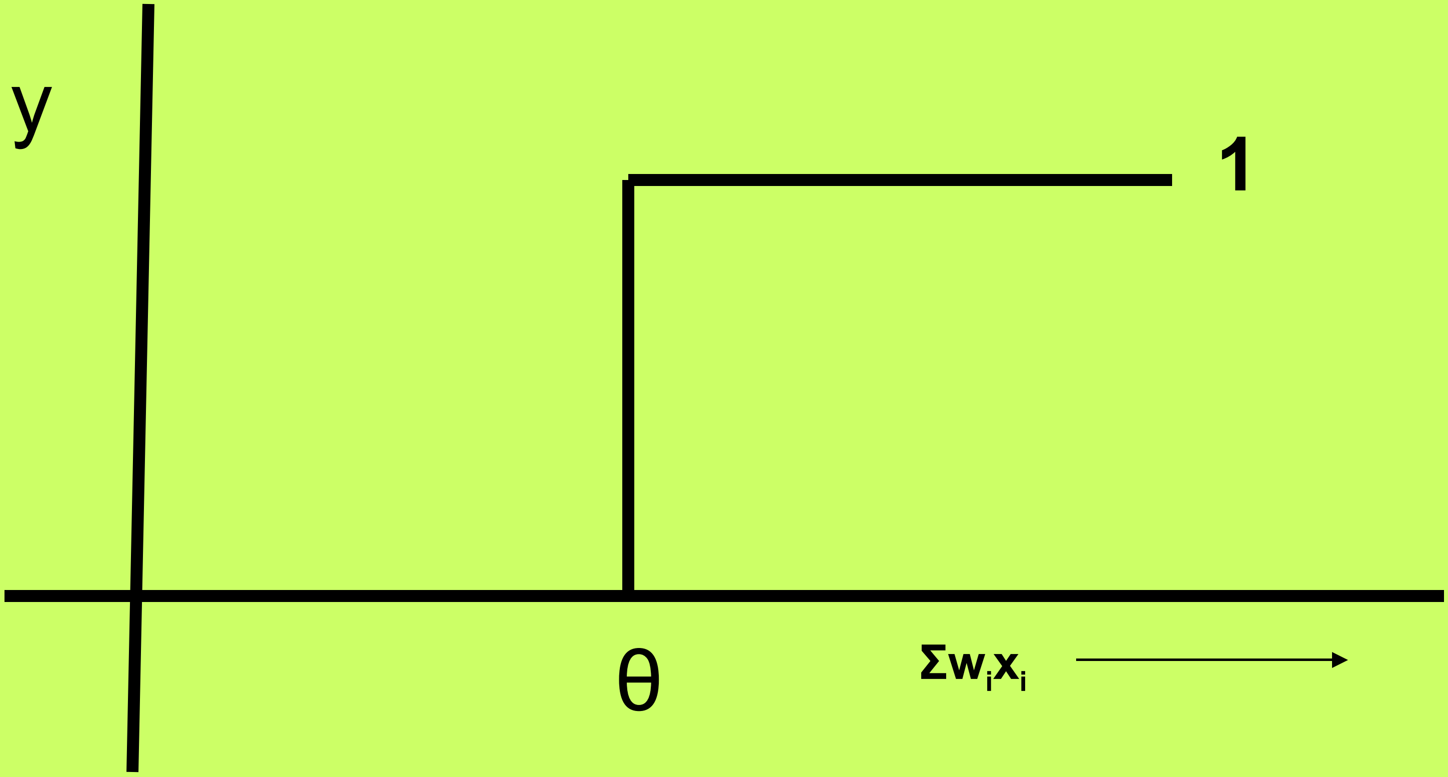
Brain computation in living beings	TM computation in computers
Pattern Recognition	Numerical Processing
Learning oriented	Programming oriented
Distributed & parallel processing	Centralized & serial processing
Content addressable	Location addressable

Perceptron

The Perceptron Model

A perceptron is a computing element with input lines having associated weights and the cell having a threshold value. The perceptron model is motivated by the biological neuron.





Step function / Threshold function

y = 1 for $\Sigma w_i x_i >= \theta$

=0 otherwise

Features of Perceptron

- Input output behavior is discontinuous and the derivative does not exist at $\Sigma w_i x_i = \theta$

- $\Sigma w_i x_i - \theta$ is the net input denoted as net

- Referred to as a linear threshold element - linearity because of **x** appearing with power **1**

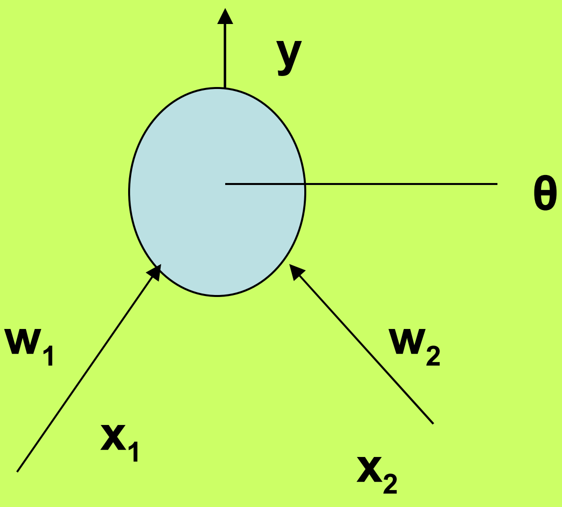
- **y= f(net)**: Relation between y and net is non-linear

Computation of Boolean functions

AND of 2 inputs

X1	x2	y
0	0	0
0	1	0
1	0	0
1	1	1

The parameter values (weights & thresholds) need to be found.



Computing parameter values

$w1 * 0 + w2 * 0 <= \theta \Rightarrow \theta >= 0$; since $y=0$

$w1 * 0 + w2 * 1 <= \theta \Rightarrow w2 <= \theta$; since $y=0$

$w1 * 1 + w2 * 0 <= \theta \Rightarrow w1 <= \theta$; since $y=0$

$w1 * 1 + w2 * 1 > \theta \Rightarrow w1 + w2 > \theta$; since $y=1$

$w1 = w2 = \; = 0.5$

satisfy these inequalities and find parameters to be used for computing AND function.

Other Boolean functions

- **OR can be computed using values of w1 = w2 = 1 and = 0.5**

- **XOR function gives rise to the following inequalities:**

  $w1 * 0 + w2 * 0 <= \theta \Rightarrow \theta >= 0$

  $w1 * 0 + w2 * 1 > \theta \Rightarrow w2 > \theta$

  $w1 * 1 + w2 * 0 > \theta \Rightarrow w1 > \theta$

  $w1 * 1 + w2 * 1 <= \theta \Rightarrow w1 + w2 <= \theta$

  No set of parameter values satisfy these inequalities.

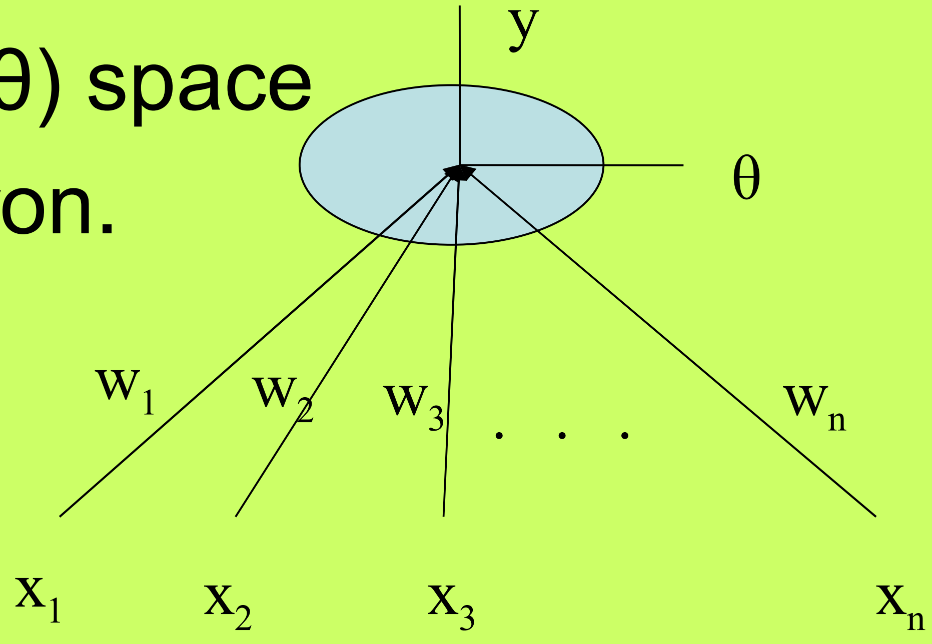
Threshold functions

n	# Boolean functions (2^{2^n})	#Threshold Functions (2^{n^2})
1	4	4
2	16	14
3	256	128
7	64K	1008

- Functions computable by perceptrons - threshold functions
- #TF becomes negligibly small for larger values of #BF.
- For n=2, all functions except XOR and XNOR are computable.

Concept of Hyper-planes

- $\sum w_i x_i = \theta$ defines a linear surface in the (W,θ) space, where $W=\langle w_1,w_2,w_3,\ldots,w_n\rangle$ is an n-dimensional vector.
- A point in this (W,θ) space defines a perceptron.



Perceptron Property

- Two perceptrons may have different parameters but same functional values.

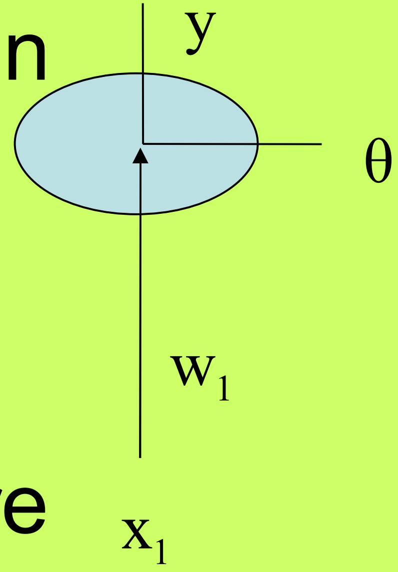
- Example of the simplest perceptron

  w.x>0 gives y=1

  w.x≤0 gives y=0

  Depending on different values of

  w and θ, four different functions are possible



Simple perceptron contd.

x	f1	f2	f3	f4
0	0	0	1	1
1	0	1	0	1

True-Function

$\theta<0$

$W<0$

0-function

$\theta\geq 0$

$w\leq 0$

Identity Function

$\theta\geq 0$

$w>0$

Complement Function

$\theta<0$

$w\leq 0$

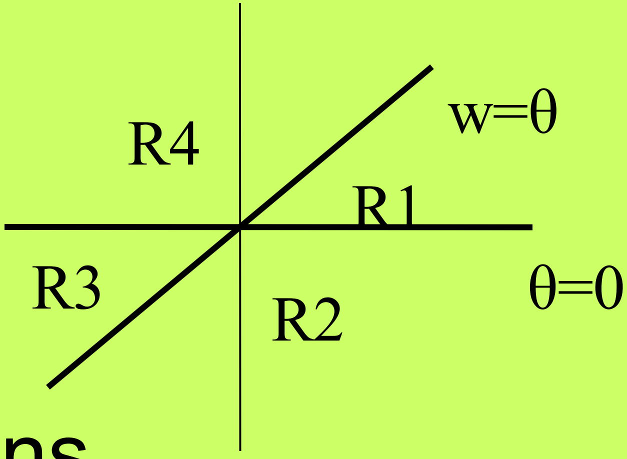
Counting the number of functions for the simplest perceptron

- For the simplest perceptron, the equation is $w.x=\theta$.

Substituting $x=0$ and $x=1$,

we get $\theta=0$ and $w=\theta$.

These two lines intersect to form four regions, which correspond to the four functions.



Fundamental Observation

- The number of TFs computable by a perceptron is equal to the number of regions produced by 2^n hyper-planes,obtained by plugging in the values $<x_1,x_2,x_3,\ldots,x_n>$ in the equation

$$\sum_{i=1}^{n} w_i x_i = \theta$$

The geometrical observation

- **Problem:** m linear surfaces called hyper-planes (each hyper-plane is of (d-1)-dim) in d-dim, then what is the max. no. of regions produced by their intersection?

  i.e. $R_{m,d} = ?$

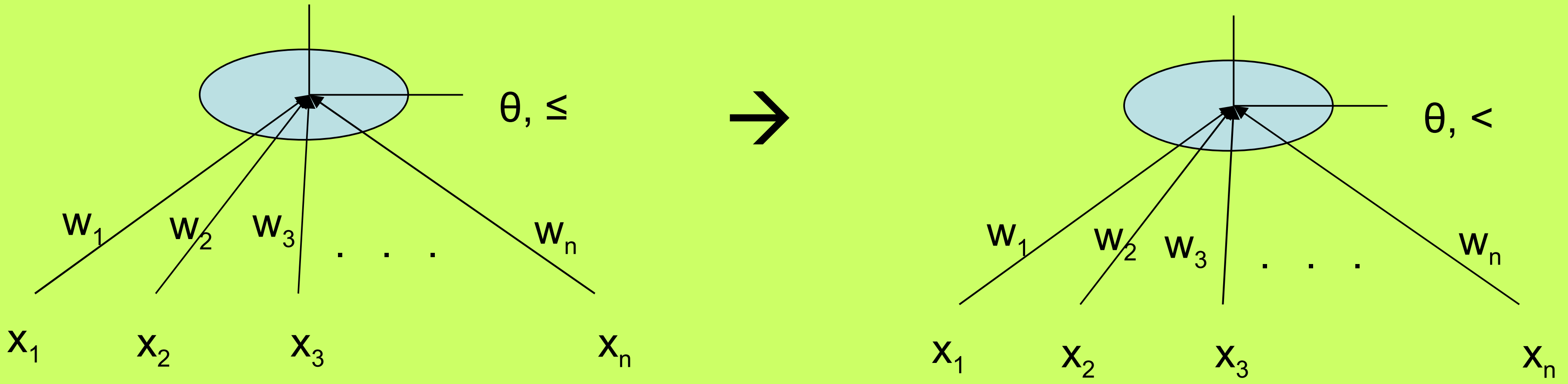
Perceptron Training Algorithm (PTA)

Preprocessing:

2. The computation law is modified to

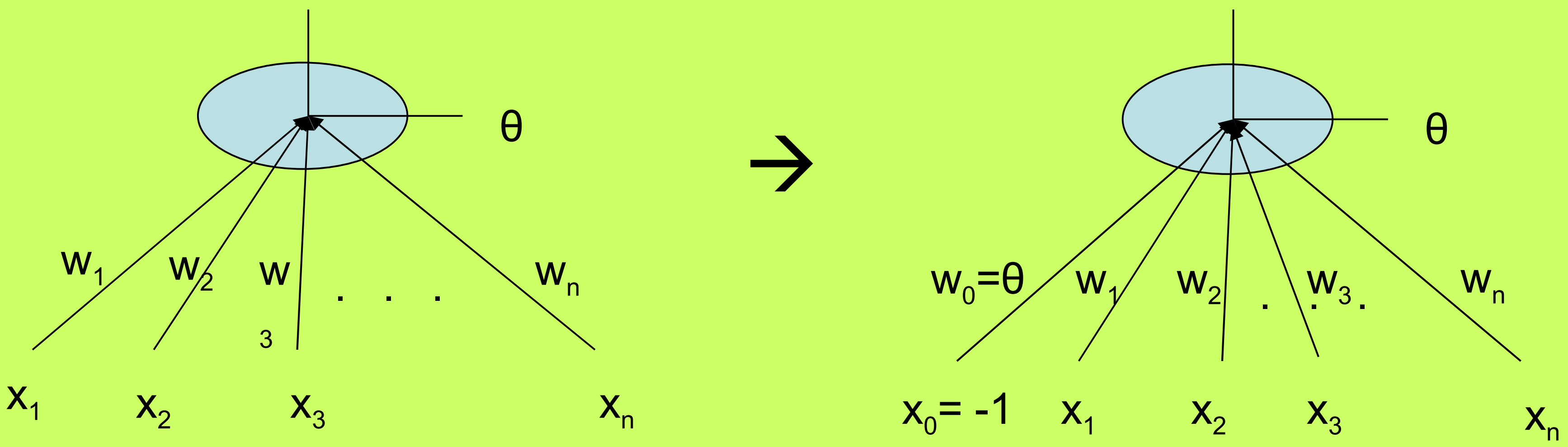
$$y = 1 \text{ if } \sum w_i x_i > \theta$$

$$y = o \text{ if } \sum w_i x_i < \theta$$



PTA – preprocessing cont…

2. Absorb θ as a weight



3. Negate all the zero-class examples

Example to demonstrate preprocessing

- **OR perceptron**

1-class <1,1> , <1,0> , <0,1>

0-class <0,0>

Augmented x vectors:-

1-class <-1,1,1> , <-1,1,0> , <-1,0,1>

0-class <-1,0,0>

Negate 0-class:- <1,0,0>

Example to demonstrate preprocessing cont..

Now the vectors are

	x_0	x_1	x_2
X_1	-1	0	1
X_2	-1	1	0
X_3	-1	1	1
X_4	1	0	0

Perceptron Training Algorithm

1. Start with a random value of w

   ex: <0,0,0…>

3. Test for $wx_i > 0$

   If the test succeeds for i=1,2,…n

   then return w

3. Modify w, $w_{next} = w_{prev} + x_{fail}$

Tracing PTA on OR-example

$w=<0,0,0>$	wx_1 fails
$w=<-1,0,1>$	wx_4 fails
$w=<0,0,1>$	wx_2 fails
$w=<-1,1,1>$	wx_1 fails
$w=<0,1,2>$	wx_4 fails
$w=<1,1,2>$	wx_2 fails
$w=<0,2,2>$	wx_4 fails
$w=<1,2,2>$	success

Assignment

- Implement the perceptron training algorithm
- Run it on the 16 Boolean Functions of 2 inputs
- Observe the behaviour
  - Take different initial values of the parameters
  - Note the number of iterations before convergence
  - Plot graphs for the functions which converge